

Acta Crystallographica Section D
**Biological
Crystallography**
ISSN 0907-4449

Structural study of the type II 3-dehydroquinase dehydratase from *Actinobacillus pleuropneumoniae*

D. Maes,^{a,*} L. A. Gonzalez-Ramirez,^b J. Lopez-Jaramillo,^b B. Yu,^{c,d} H. De Bondt,^d I. Zegers,^a E. Afonina,^{c,d} J. M. Garcia-Ruiz^b and S. Gulnik^{c,d}

^aULTR, Vrije Universiteit Brussel, Vlaams Interuniversitair Instituut voor Biotechnologie (VIB), Pleinlaan 2, 1050 Brussels, Belgium, ^bLaboratorio de Estudios Cristalográficos, Instituto Andaluz de Ciencias de la Tierra, CSIC–Universidad de Granada, Av. Fuentenueva s/n, 18002 Granada, Spain, ^cSequoia Pharmaceuticals Inc., 401 Professional Drive, Suite 100, Gaithersburg, MD 20879, USA, and ^dTibotec BVBA, Generaal De Wittelaan 11B 3, B-2800 Mechelen, Belgium

Correspondence e-mail:
dominique.maes@vub.ac.be

The structure of the type II dehydroquinase dehydratase (DHQase) from *Actinobacillus pleuropneumoniae*, the third enzyme of the shikimate pathway, has been determined. Crystals diffracting to 1.7 Å were obtained in space and on earth using the counter-diffusion technique. The structure was solved using molecular replacement and refined to high resolution. The overall structure of the dodecameric enzyme is described and compared with structures of DHQases from other bacteria. DHQases contain a flexible loop that presumably closes over the active site upon substrate binding. The enzyme can exist in an open or closed conformation. The present structure displays the open conformation, with a sulfate anion bound in the active site. The availability of this structure opens a route to structure-based antibiotics targeting this pathogenic bacterium.

1. Introduction

This paper is concerned with the type II dehydroquinase (DHQase) from *Actinobacillus pleuropneumoniae* (referred to here as Ap-DHQase). DHQase catalyses the dehydration of 3-dehydroquinase (DHQ) to 3-dehydroshikimate (Fig. 1), a reaction common to two metabolic pathways that are absent in animals: (i) the biosynthetic shikimate pathway, which is used for synthesis of aromatic compounds (present in microorganisms and plants), and (ii) the catabolic quinate pathway, which enables the use of quinate as a carbon and energy source (present in fungi and some microorganisms). DHQases have been found to fall into two main classes: type I and type II. Representative structures of DHQases of both types have recently been solved (Gourley *et al.*, 1999). The two classes of enzymes operate by entirely different mechanisms; their sequences are unrelated and they have different folds and subunit organization. The type I enzyme is a dimer of eight-stranded α/β -barrels, while the type II enzyme has a dodecameric quaternary structure. The existence of two distinct types of DHQase in microorganisms offers a unique opportunity to selectively inhibit some pathogenic microorganisms using the type II enzyme as a drug target (Coggins *et al.*, 2003).

Actinobacillus (formerly *Haemophilus*) *pleuropneumoniae* is an aetiological agent of porcine pleuropneumonia, a major respiratory disease in pigs (for recent reviews, see Bosse *et al.*, 2002; Lalonde *et al.*, 1994; Rycroft & Garside, 2000, and references therein). The wide spread of the disease causes severe economic losses to the pig industry worldwide. Infection is transmitted by contact or aerosol and the morbidity in susceptible groups can approach 100%. Infected animals develop acute fibrinous pneumonia, which leads to death or chronic lung lesions and reduced growth rates. Commercially available vaccines, which are generally composed of chemi-

Received 21 October 2003
Accepted 23 December 2003

PDB Reference: type II
3-dehydroquinase dehydratase, 1uqr, r1uqrsf.

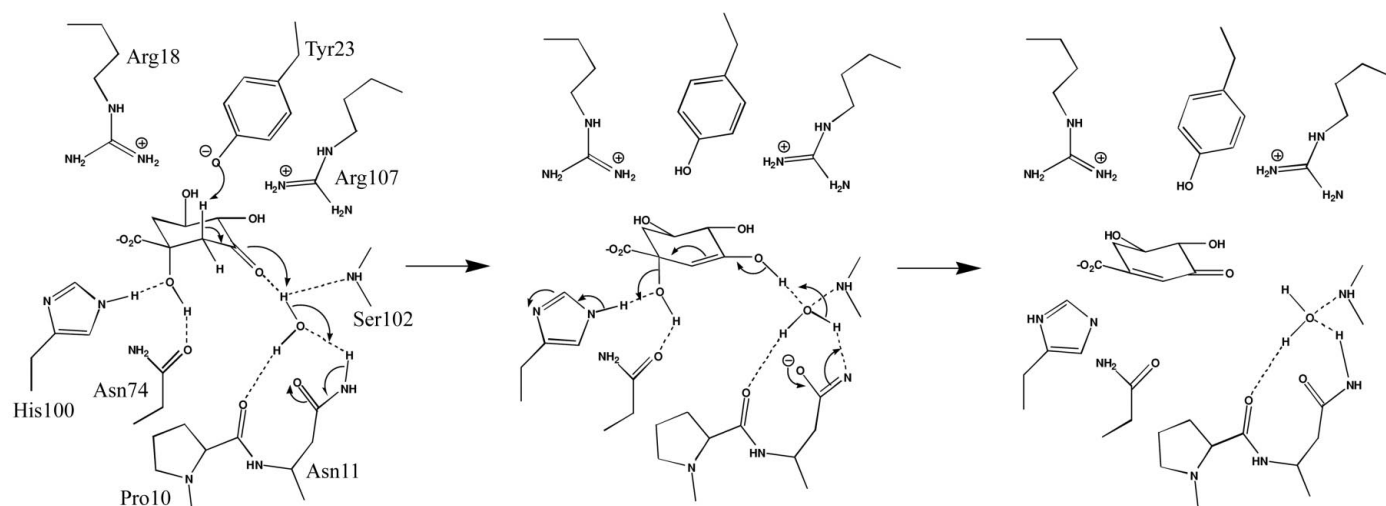


Figure 1

Summary of the reaction mechanism of type II dehydroquinases (Roszak *et al.*, 2002).

cally inactivated bacteria combined with oil adjuvants, have failed to prevent the disease. Moreover, whole-cell bacterins and surface-protein extracts often contain immunosuppressive components, which make pigs more susceptible to infection. Furthermore, these vaccines may reduce mortality but do not reduce the number of chronic carriers in a herd. Several antibiotics have been shown to be effective against *A. pleuropneumoniae*. As the overuse of animal drugs could spur the emergence of superbacteria in meat that can infect consumers and resist drugs used to treat humans, there is a clear need for effective and more specific antibiotics. In this respect, type II dehydroquinase (DHQase) may represent a good target for such increased-specificity drugs, since most common meat-contaminating bacteria that affect humans, such as *Listeria monocytogenes*, *Escherichia coli* and various *Salmonella* species, have structurally and mechanistically different type of DHQase (type I DHQase) and therefore will not develop resistance to type II inhibitors.

At present, the structures of type II DHQases from *Mycobacterium tuberculosis* (Mt-DHQase; Gourley *et al.*, 1999), *Bacillus subtilis* (Bs-DHQase), *Streptomyces coelicolor* (Sc-DHQase; Roszak *et al.*, 2002) and *Helicobacter pylori* (Hp-DHQase; Lee *et al.*, 2003) have been solved and are available in the PDB. For Mt-DHQase the unliganded and sulfate-liganded structures have been solved, while the Bs-DHQase structure contains glycerol molecules, the Hp-DHQase structure contains the substrate dehydroquinone, and Sc-DHQase structures are available that are unliganded and phosphate-liganded as well as liganded with the transition-state analogue 2,3-anhydroquinic acid and with the product dehydroshikimate as well as tartrate and glycerol for the inactive R23A (Sc-DHQase numbering) mutated enzyme.

In this paper, we present the high-resolution (1.7 Å) crystal structure of the type II DHQase from *A. pleuropneumoniae* (Ap-DHQase) containing sulfate anions. Ap-DHQase was crystallized by the counter-diffusion technique both in space under microgravity conditions and on earth. The use of the counter-diffusion technique increased the diffraction limit

from the crystals from 3 to 1.7 Å. The structure of Ap-DHQase shows that it has a flavodoxin-like fold and a dodecameric quaternary organization with tetrahedral symmetry, similar to the other type II DHQase structures. The active-site region contains a flexible loop. This flexible loop, which is not visible in most of the other available unliganded structures of type II DHQases, could be traced in several subunits in the present structure. Different conformations were found, showing its flexibility. This loop domain, which includes the essential catalytic residues Arg18 and Tyr23 (Ap-DHQase numbering), forms a lid over the active site in the liganded structures.

Unless specified otherwise, Ap-DHQase numbering is used.

2. Material and methods

2.1. Cloning, expression and purification

Genomic DNA from *A. pleuropneumoniae* was obtained from ATCC. The Ap-DHQase gene (Lalonde *et al.*, 1994) was amplified by PCR using *A. pleuropneumoniae* genomic DNA as template and the following primers: sense, atatatCATATGAAAAAATTCTTCTGTAAACG; antisense, atatatGGATCCTTAGCCATTCATCATTTCTC. The PCR fragment was inserted into pET21a (Novagen) via *Nde*I and *Bam*HI sites. The sequence of the resulting construct, pET21DHQ *pne*, was confirmed by DNA sequencing. pET21DHQ *pne* was transformed into BL21(DE3) *E. coli* cells (Novagen, Milwaukee, WI, USA). A single colony was inoculated into 50 ml LB medium containing 0.1 mg ml⁻¹ ampicillin. The culture was grown overnight at 310 K with shaking. The cells from 10 ml aliquots were harvested by centrifugation, resuspended in 10 ml fresh LB/ampicillin and used to inoculate 1000 ml LB/ampicillin medium. The culture was maintained at 310 K with shaking until the OD at a wavelength of 600 nm reached 0.7. At this stage, the culture was induced by IPTG at a final concentration of 1 mM. After 3 h of further growth, the cells were harvested by centrifuga-

tion and stored at 193 K. The cell paste obtained from 1 l of culture was resuspended in 20 ml ice-cold 50 mM Tris–HCl pH 7.5 and sonicated twice for 2 min with a 10 min rest on ice in between. The crude cell extract was centrifuged at 47 000g at 277 K for 60 min and the supernatant was incubated at 348 K for 15 min with occasional stirring. The solution was then cooled on ice for 20 min and heat-denatured proteins were removed by centrifugation at 47 000g for 30 min. The proteins in the supernatant were precipitated with 65% saturated ammonium sulfate. The precipitate was collected by centrifugation at 26 000g for 30 min, redissolved in 5 ml 50 mM Tris–HCl buffer pH 7.5 and chromatographed through a 2.6 × 60 cm Sephacryl S-200 column (Amersham) using an isocratic elution in the same buffer. Fractions containing Ap-DHQase were pooled and applied onto a 1.0 × 10 cm MonoQ column (Amersham). Pure protein was eluted using a 0–1 M linear gradient of NaCl in 50 mM Tris–HCl buffer pH 7.5. The total yield of purified protein was 60 mg per litre of *E. coli* culture. The protein concentration was estimated based on the absorbance at 280 nm using a calculated extinction coefficient of 0.45 for 1 mg ml^{−1} protein solution.

2.2. Crystallization

Crystals were grown by the gel-acupuncture method (Garcia-Ruiz *et al.*, 1998) implemented in the Granada Crystallization Box (GCB; Garcia-Ruiz *et al.*, 2002) both on earth and in space. The space experiments were performed in the Granada Crystallization Facility in the context of the Andromeda mission.

The GCB was filled to 2.7 cm in height with a gel consisting of 1% (w/v) agarose in 100 mM Tris–HCl pH 7.4 and 0.03% (w/v) sodium azide. Six X-ray capillaries with internal diameters of 0.2, 0.3, 0.5, 0.6, 0.7 and 1 mm were filled with a solution of Ap-DHQase at a concentration of 20 mg ml^{−1} in 100 mM Tris–HCl pH 7.4. The capillaries were punctured 1 cm into the gel layer to delay crystal growth for 72 h in order to be sure that the experiment was orbiting the Earth. Finally, the precipitating agent, consisting of 3.5 M ammonium sulfate in 100 mM Tris–HCl pH 7.4 and 0.03% (w/v) sodium azide, was poured onto the gel to a total height of 8 cm. The GCB flew to the International Space Station (ISS) in a Russian Progress M-45 logistics vehicle launched on 21 August 2001 and returned to Earth on board a Russian Soyuz TM-33 manned transport vehicle launched on 21 October that landed on 31 October 2001. The total length of this Andromeda Mission was therefore 72 d. A set of control experiments on ground were prepared using exactly the same conditions except that the protein solution was gelled inside the capillary with 0.1% agarose to prevent convective motion (Garcia-Ruiz, Novella *et al.*, 2001). As shown in Fig. 2, both ground and space experiments yielded numerous crystals distributed along the capillaries with the typical counter-diffusion crystal distribution along the capillary, with crystals becoming larger towards the top of the capillary.

2.3. Data collection and data processing

Two data sets were collected, one from a crystal grown in space and one from a crystal grown on earth, both crystals having the same size, morphology and habit. The crystals were removed from the capillaries and soaked in a cryobuffer consisting of 100 mM Tris–HCl pH 7.4, 2.5 M ammonium sulfate and 20% (w/v) glycerol. Data were collected at Hamburg EMBL synchrotron outstation (beamline BW7b) at 100 K in dose mode. The images were recorded on a MAR 345 image-plate detector placed at a distance of 270 mm with an oscillation angle of 0.3°. Data sets were scaled with *XSCALE*. Data reduction was carried out independently for each data set with the program package *XDS* (Kabsch, 1993) and neither overloads nor incomplete reflections (*i.e.* those for which the estimated intensity after profile fitting was less than 75% of the observed intensity) were included in the processing. Finally, common reflections from each data set were merged together in order to solve the structure.

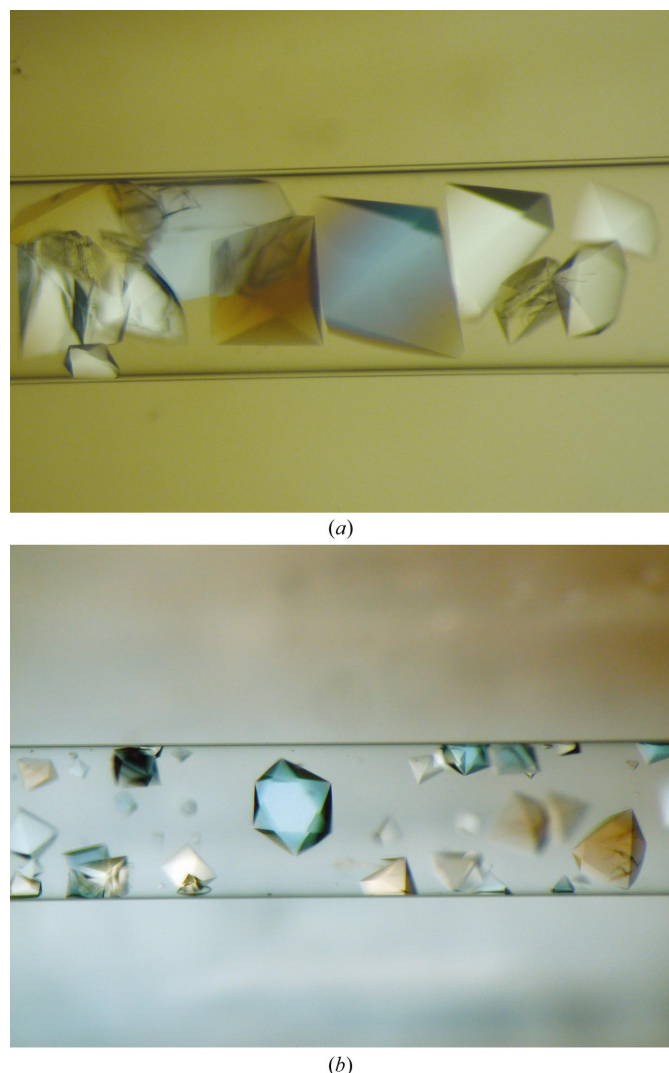


Figure 2
Crystals of DHQase grown (a) in space inside a 0.7 mm capillary and (b) on earth inside a 1.0 mm capillary.

2.4. Molecular replacement, model building, refinement and structure analysis

Molecular replacement was performed with the program *AMoRe* (Navaza, 1994, 2001) using the 1.7 Å data set from a ground-control crystal. For computational reasons, half of the dodecamer, consisting of chains *A* to *F*, of Sc-DHQase (PDB code 1d0i) was used as a search model. Because of the symmetry of the molecule, several solutions were found. Two non-overlapping solutions were chosen and this resulted in a model with an *R* factor of 48.9% after rigid-body refinement.

Manual rebuilding and computational refinement were performed using the programs *TURBO-FRODO* (Roussel & Cambillau, 1991) and *CNS* (Brünger *et al.*, 1998) running on an SGI O2 workstation. The parameters for ideal protein geometry of Engh & Huber (1991) were used for the refinement. Tris and sulfate molecules were identified based on the shape of the electron density and included in the model. Water peaks were assigned using the standard protocol as implemented in *CNS*. Briefly, peaks greater than 3σ in an $F_{\text{obs}} - F_{\text{calc}}$ electron-density map with at least one potential hydrogen bond were added to the model and continuously updated during the refinement (*B* factors were not allowed to be greater than 62 Å^2). Both *R* and R_{free} (5% of reflections in the test set) were used to monitor the progress of structural refinement (Brünger, 1992). The stereochemical quality of the refined structure was checked with the program *PROCHECK* (Laskowski *et al.*, 1993).

Superpositions were carried out with *LSQMAN* (Kleywegt & Jones, 1994) and *TURBO-FRODO* (Roussel & Cambillau, 1991). All C^α atoms except those in the terminal residues, in the residues in the flexible loop and in a small region around insertion 92 were superimposed (residues 5–17, 25–91, 97–142). Figures were drawn with *MOLSCRIPT* (Kraulis, 1991) and *Raster3D* (Merritt & Bacon, 1997). *CLUSTALW* (Aiyar, 2000) was used for multiple sequence alignment. For the comparison of the active sites, a short stretch of structurally well conserved residues (HL/IS/TN motif; residues 100–103; Evans *et al.*, 2002) involved in substrate recognition was superimposed.

3. Results and discussion

3.1. Protein production and purification

The Ap-DHQase gene was cloned from bacterial genomic DNA and inserted into the pET 21 vector for *E. coli* expression using conventional molecular biology techniques. The protein was overexpressed in soluble form and purified to apparent homogeneity using a simple purification procedure that takes advantage of the excellent thermal stability and high molecular weight of the dodecameric enzyme. Ap-DHQase remained folded at 348 K, while most of *E. coli* proteins underwent heat denaturation and were removed by centrifugation. Further purification was achieved by size-exclusion chromatography of the 205 kDa dodecamer. Ion-exchange chromatography on Mono-Q allowed separation of a significant amount of material absorbing at 254 nm, probably

Table 1

Data-collection and processing statistics.

Values in parentheses are for the highest resolution shell.

	Space	Ground
Beamline (DESY)	BW7b	BW7b
Wavelength (Å)	0.8463	0.8463
Distance to detector (mm)	270	270
Oscillation angle (°)	0.3	0.3
Images collected	100	416
No. crystals	1	1
Cryoconditions	Yes	Yes
Unit-cell parameters (Å, °)	$a = 129.09$, $b = 131.33$, $c = 161.62$, $\alpha = \beta = \gamma = 90$	$a = 128.96$, $b = 131.35$, $c = 160.84$, $\alpha = \beta = \gamma = 90$
Space group	$P2_12_12_1$	$P2_12_12_1$
Resolution range (Å)	20.00–1.71 (1.80–1.71)	20.00–1.71 (1.80–1.71)
Unique reflection	210933 (27619)	288020 (39797)
Completeness (%)	71.5 (66.2)	98.2 (96.0)
R_{sym} (%)	3.3 (30.9)	3.5 (22.4)
Redundancy	1.64	5.12
<i>B</i> factor from Wilson plot (Å ²)	26.0	27.2
$I/\sigma(I)$	12.67 (2.90)	27.78 (7.30)
Completeness at $I/\sigma(I) \geq 2$ (%)	60 (37)	93 (79)

bound nucleic acids, and resulted in more than 99% pure protein as judged by SDS–PAGE analysis.

3.2. Crystallization by counter-diffusion

Iterative trials to crystallize DHQase by the vapour-diffusion technique never yielded crystals that diffracted beyond 3 Å using in-house diffraction facilities. Therefore, out-of-equilibrium counter-diffusion experiments were designed in order to improve crystal quality. The principle of counter-diffusion experiments involves the diffusion of a precipitant (ammonium sulfate in this case) into a long capillary containing the protein solution. A layer of gelled buffer initially separates the precipitant solution and the protein solution. The protein molecules diffuse out of the capillary much more slowly than the salt molecules. The coupling of the diffusing salt and protein molecules with the precipitation of the protein provokes a wave of supersaturation moving over the capillary. The crest of the wave lowers as the diffusion process proceeds and the precipitant reaches the upper part of the capillary. Thus, different parts of the capillary ‘see’ different crystallization conditions. In the capillary, a continuum of crystallization conditions is created from high supersaturation when the protein and salt first meet at the lower part of the capillary to low supersaturation at the upper part of the capillary. Thus, the out-of-equilibrium counter-diffusion method self-searches for the best crystallization conditions and the protein crystallizes in the region with supersaturation profiles that are best suited (Garcia-Ruiz, Otalora *et al.*, 2001). The starting conditions for counter-diffusion were based on previous vapour-diffusion conditions. Fig. 2 shows ground- and space-grown crystals. Both are faceted and display good optical quality. Owing to a lack of synchrotron time, X-ray crystallographic data were only collected from two crystals: one crystal grown in space and one

crystal grown on the ground. The results of the data analysis are summarized in Table 1. The use of counter-diffusion improved the resolution from 3.0 Å for the crystals grown in hanging drops to 1.7 Å. This result cannot be considered conclusive because of three reasons: (i) a proper comparative analysis of crystal quality between space and earth-grown crystals grown by counter-diffusion technique needs to be performed with a representative set of crystals because, as discussed above, the quality of the crystals varies as a function of their location along the capillary, (ii) although both data sets were collected with the same parameters (*e.g.* beam, exposure, distance to the detector, oscillation angle), direct comparison is not straightforward because the crystals were not oriented and the oscillation range of each data set is different and (iii) the thermal history of the space experiments was unknown.

3.3. The Ap-DHQase molecular structure

The crystal structure of Ap-DHQase was determined by molecular replacement using the Sc-DHQase structure (Roszak *et al.*, 2002) as a phasing model. The final model has good stereochemistry. Ap-DHQase exists as a homo-dodecamer: four trimers are arranged in a tetrahedral manner (23 symmetry; Fig. 3*a*). Each subunit consists of a five-stranded parallel β -sheet core flanked by four α -helices (Fig. 3*b*). The interactions between the trimeric units are less extensive than those within the trimers. On forming the trimer, each monomer buries 1329 Å² of solvent-accessible surface (17% of the total surface). During assembly of the trimers, an additional surface of 813 Å² (10% of the total surface) is buried.

The asymmetric unit consists of the complete dodecamer. For most monomers there is only weak density for the C-terminal helix (residues 146–153) and for residues 18–25. This flexible loop forms a lid that closes over the active site on substrate binding. After several rounds of refinement and incorporation of water molecules, we were able to trace the flexible loop for four subunits and the C-terminal helix for five subunits. The side chain of Cys125 adopts three different conformations in all subunits. In one of its conformations Cys125 makes a hydrophobic intertrimeric contact with Val123. The 12 monomers are highly similar, with r.m.s. deviations ranging between 0.13 and 0.78 Å for the superposition of the main-chain C α atoms of all the possible pairs of monomers. The largest differences between monomers are located at the flexible loop and at the C-terminal helix. The refined model includes 2266 structured water molecules, four Tris molecules and 31 sulfate molecules. Refinement parameters and statistics are given in Table 2.

3.4. The active site and the flexible loop

The type II DHQases catalyze *trans*-dehydration *via* an enol intermediate (Fig. 1). The active site is composed of residues Pro10, Asn11, Arg18, Tyr23, Asn74, Ala77, Ser102 and Arg107 (Roszak *et al.*, 2002) and is located at the C-terminal end of the five-stranded parallel β -sheet (Fig. 3*b*). In the

Table 2

Refinement parameters and statistics for the final model.

Protein atoms	13650
Sulfate atoms	31
Tris atoms	4
Solvent atoms	2266
Resolution range (Å)	20–1.71
<i>R</i> factor (%)	17.8
<i>R</i> _{free} (%)	19.9
R.m.s. bond-length deviations (Å)	0.006
R.m.s. bond-angle deviations (°)	1.3
Average <i>B</i> factors (Å ²)	
Protein atoms	23.7
Sulfate atoms	38.6
Tris atoms	17.3
Solvent atoms	40.6
Ramachandran plot	
Most favoured regions (%)	92.6
Additional allowed regions (%)	6.8
Generously allowed regions (%)	0.5
Disallowed regions (%)	0.0

present structure, the 12 active sites are occupied by several structured water molecules and a sulfate ion bound in the carboxylate-recognition region making hydrogen bonds with the main-chain N atoms of Ile101 and Ser102 and with Asn74 ND2, His100 ND1 and Ser102 OG. The active-site residue Arg18 is located at the N-terminal region of the flexible loop. The deprotonated form of Tyr23, belonging to the loop, has been proposed to act as a general base, abstracting a proton from C2 of the substrate. Arg107 and to a lesser extent Arg18 could lower the p*K*_a of Tyr23. In none of the monomers does Tyr23 take up the position it should have for catalysis. It is more than 5 Å away from C2 of a modelled substrate. Thus, the different monomers in the present structure represent different conformations of the open form of Ap-DHQase. The loop residues are rather poorly defined (high temperature factors), showing their flexibility/mobility. Nevertheless, the complete flexible loop was traced in four subunits (*A*, *D*, *J*, *K*). The main chain follows the same trace in three subunits (*D*, *J*, *K*). For subunit *A* a completely different conformation was found for this loop. This is because of the close crystallographic packing contact in which this loop is involved: direct hydrogen bonds are formed with His34, Gln37, Ser38, Glu42, Lys130 and Asp133 of chain *C* as well as with Asp137 of chain *L*. Comparison of the two conformations of the loop found in the present structure and the typical closed conformation (*e.g.* torsion angles) shows that this loop motion cannot be seen as a hinge motion.

3.5. The non-specific ligands

A large number of ligands (four Tris molecules and 31 sulfate molecules) that were present in the crystallization solutions were identified in the electron density.

A Tris molecule is located on the threefold axis of each trimer. Its three symmetric O atoms make close contacts with the O atoms of the side chains of Glu54 and Thr81 from the three subunits forming the trimer.

31 sulfate molecules were clearly visible in the electron density. They can be divided in five different classes as follows (Fig. 3*b*).

(i) A sulfate (yellow) is located on the threefold axis of each trimer, 8.5 Å from the corresponding Tris molecule. The sulfate is hydrogen bonded to the main-chain N atoms of residues Glu54 and Glu55 of the monomers forming the trimer.

(ii) In all 12 subunits a sulfate ion (green) is located in the active site and is described in §3.4.

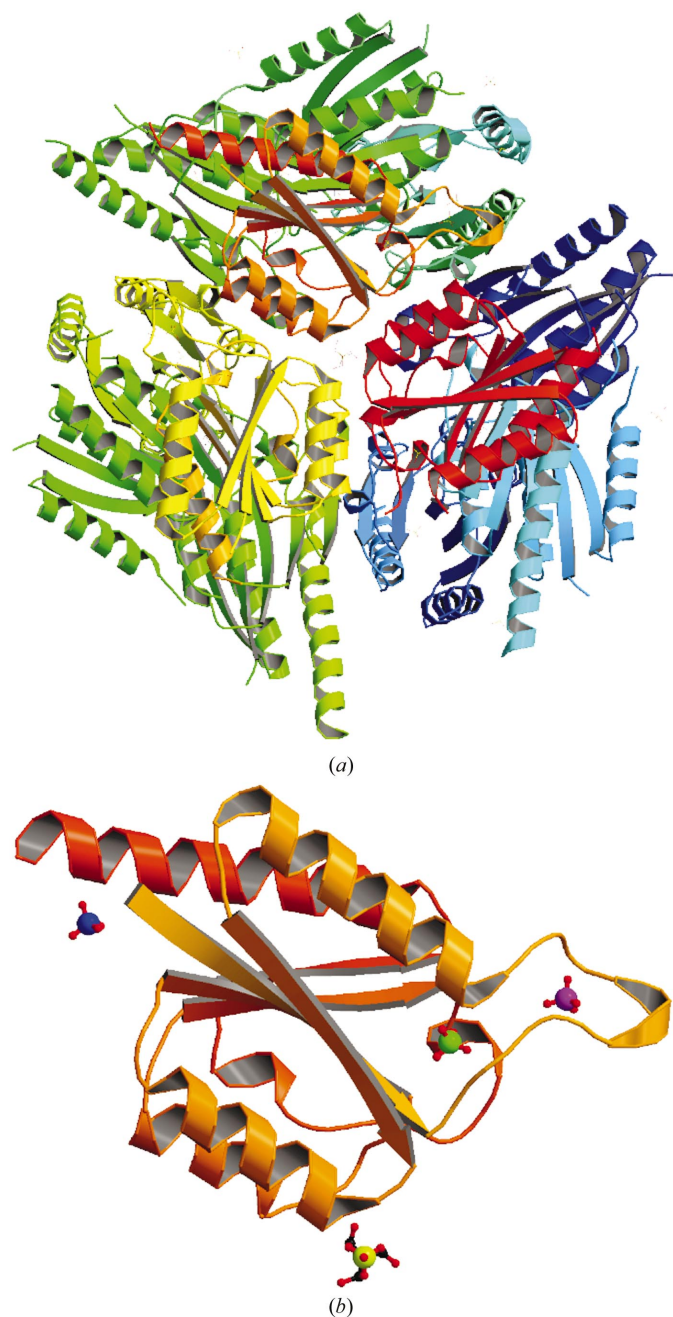


Figure 3
(*a*) Ribbon diagram of the dodecameric quaternary structure of Ap-DHQase. (*b*) Ribbon diagram of the orange-coloured subunit in the same orientation as in (*a*). The sulfates and the Tris molecule are shown in ball-and-stick representation.

(iii) In seven of the subunits a sulfate ion is located at the surface of the protein, making a contact with Thr27 OG1. The subunits that do not have this sulfate have packing contacts in this region. In the *F* chain this sulfate mediates a packing contact with the C-terminal region of the *K* chain.

(iv) In seven subunits a sulfate ion (violet) was modelled into the density at the entrance of the active site in the neighbourhood of the flexible loop. This sulfate always occupies the same position in the same conformation. We observe that in subunits in which we were able to trace the flexible loop (*e.g.* *K*), this sulfate is hydrogen bonded to the main-chain N atoms of Gly16, Lys17, Ile22, Tyr23 and Gly24 (five main-chain hydrogen bonds) and to Arg18 NH2 and His21 ND1. The latter contact is only present in chain *K*. In the other chains this histidine (if ordered in the structure) is in a different conformation. The presence of this sulfate, with its extensive main-chain contacts, probably results in a more rigid active-site loop. The absence of this sulfate in chain *A* is associated with a different conformation of the flexible loop.

(v) One sulfate ion (blue) is located in the N-terminal part of chain *K* at a distance of 3.7 Å from the N atom of Met1. This sulfate makes packing contacts with the side chains of Tyr48 and Gln36 of a symmetry-equivalent molecule and has a low temperature factor.

As expected, the first two classes of sulfates, which form a structured part of the molecule, have low temperature factors ($<30 \text{ Å}^2$). In contrast, sulfates from the third class, which are located at the surface of the protein, and from the fourth class, which are located in a flexible region of the molecule, have high temperature factors ($>50 \text{ Å}^2$).

3.6. Structural comparison with other DHQases

3.6.1. Molecular structure. The structures of Ap-DHQase, Sc-DHQase, Bs-DHQase, Hp-DHQase and Mt-DHQase were compared. The alignment of the amino-acid sequences is given in Fig. 4. As expected, the dodecameric quaternary structure organization as well as the polypeptide fold are strikingly similar for the four DHQases. Secondary-structure elements are highly conserved. Superposition of the C^α coordinates of Sc-DHQase (PDB code 1gu0), Bs-DHQase (PDB code 1gq0), Hp-DHQase (PDB code 1j2y) and Mt-DHQase (PDB code 1h05) on chain *K* of Ap-DHQase gives r.m.s.d.s of 1.2, 1.4, 1.4 and 1.2 Å, respectively. The strands forming the core of the enzyme have a lower r.m.s.d. than the helices. Superposition of the different structures shows that the largest differences are located in the flexible loop and at both termini of the polypeptide chain. Significant r.m.s. differences are also found in the loops at the C-terminal end of helix 1, helix 2 and helix 3, the last loop having an insertion of one residue in Sc-DHQase (Fig. 5). These three loops are located opposite the active site of the molecule.

Of the two intermonomeric salt bridges within the trimer (Arg86–Glu108 and Arg111–Asp87), the latter is present in all known type II DHQase structures. This interaction contributes to the closing of the active-site cleft. The first salt bridge is not found in Sc-DHQase and Hp-DHQase.

	1	β 1	10	20	30	α 1	40	β 2	50	
Ap-DHQase	----	MKKILLNGP	NLNMLGKREP	HIYGSQTLSD	IEQHLQSSAQ	AQG--YELDYFQ	ANGEE			
Sc-DHQase	PRSLA	NAPIMILNGP	NLNLLGQRQP	EIYGSQTLAD	VEALCVKAAA	AHG--GTVDFRQ	SNHEG			
Mt-DHQase	-----	ELIVNVINGP	NLGRIGRRREP	AVYGGTTTDE	LVALIEREAA	ELG--LKAVVRQ	SDSEA			
Hp-DHQase		MKILVIQGP	NLNMLGHRDP	RLYGMVTLDQ	IHEIMQTFVK	QGNLDVELEFFQ	TNFEF			
Bs-DHQase	-----	MPHFLILNGP	NVNRLGSRREP	EVFGRQTLTD	IETDLFQFAE	ALH--IQLTFFQ	SNHEG			
		*** *	** *	*** *	*** *	*** *	*** *	*** *	*** *	*** *
				fff	ffff					
	α 2	60	70	β 3	80	α 3	90	β 4	100	110
Ap-DHQase	SLINR	IHQAFQN--TDF	IIINPGAFTH	TSVAIRDALL	AVS-IPFIEVH	LSNVHAREPF	RHHS			
Sc-DHQase	ELVDW	IHEARLNH--CG	IVINPAAYSH	TSVAILDALN	TCDGLPVVEVH	ISNIHQREPF	RHHS			
Mt-DHQase	QLLDW	IHQADAA--EP	VILNAGGLTH	TSVALRDACA	ELS-APLIEVH	ISNVHAREEF	RRHS			
Hp-DHQase	EIIDK	IQESVGS-DYEG	IIINPGAFSH	TSIAIADAIM	LAG-KPVIEWH	LTNIQAREEF	RKNS			
Bs-DHQase	DLIDA	IHEAEEQ--YSG	IVLNPGALSH	YSYAIRDAVS	SIS-LPVVEVH	LSNLYAREEF	RHQS			
	*	** *	*	*	** *	*	***	**	**	** *
	120	β 5	130	α 4	140	150	160			
Ap-DHQase	YLSOVA	KGVICGLGAK	GYDYALDFAI	SELQKIQLGE	MMNG-----	-----	-----			
Sc-DHQase	YVSQRA	DGVVAGCGVQ	GYVFGVERIA	ALAGAGSARA	-----	-----	-----			
Mt-DHQase	YLSPIA	TGVIVGLGIQ	GYLLALRYLA	EHVGT-----	-----	-----	-----			
Hp-DHQase	YTGAAC	GGVIMFGFPL	GYNMALMAMV	NILAEMKAFQ	EAQKNNPNNP	INNQK	-----			
Bs-DHQase	VIAPVA	KGQIVGLGAE	GYKLAVRYLL	SQQGGESR--	-----	-----	-----			
	*	** *	** *	** *	** *	** *	** *			

Figure 4

Structural sequence alignment of Ap-DHQase, Sc-DHQase, Mt-DHQase, Hp-DHQase and Bs-DHQase. Active-site residues are in bold; an asterisk indicates conserved residues and an 'f' residues belonging to the flexible loop. Residues belonging to secondary-structure elements are coloured blue for β -strands and red for α -helices. Assignment was performed with *DSSP* and by visual inspection.



Figure 5

Superposition of the backbone of chain A (blue) and chain K (yellow) from Ap-DHQase, of the two conformations of the unliganded Bs-DHQase (chain A, brown; chain K, purple; PDB code 1gq0), of the liganded Hp-DHQase (red; PDB code 1j2y) and of the unliganded (green; PDB code 1gu0) and liganded (light blue; PDB code 1gu1) Sc-DHQase. The active-site sulfate of Ap-DHQase is shown in ball-and-stick representation.

3.6.2. The flexible loop. The superposition of the different chains reveals six different traces for the flexible loop (Fig. 5); in the three liganded Sc-DHQase structures (with phosphate, with a transition-state analogue and with a product in the

R23A mutant; Sc-DHQase numbering, equivalent to Arg18 in Ap-DHQase) this loop is in the closed conformation. In the unliganded Sc-DHQase structure this loop is in an open conformation. In Bs-DHQase this loop adopts two different conformations: one conformation is the closed conformation, identical to the closed conformation of the Sc-DHQase structure, while the second conformation arises from packing contacts. Binding of glycerol is sufficient to close the loop, but the contacts the flexible loop makes are weaker than the packing contacts. In the Hp-DHQase structure, which possesses a substrate analogue in the active site, the loop is in the closed conformation. In the available structures of Mt-DHQase this loop is disordered and no coordinates are available. In the present Ap-DHQase

this loop is also in two conformations. In the first conformation it is partially similar to the closed conformation. The second conformation is an open conformation, making packing contacts. In contrast, in Sc-DHQase the flexible loop does not have an ordered helical lid domain and does not undergo a hinge movement in the present structure and in Bs-DHQase. Although this loop, which is involved in catalysis, shows large conformational flexibility, it is clear from this analysis that the closed form is identical for all DHQase structures; the open form has preferred conformations which are not similar in the DHQase structures.

3.6.3. The active site. The active-site residues are conserved in all known structures of type II DHQase and superimpose very closely. In Bs-DHQase Tyr23 is replaced by a phenylalanine. Although no activity data have been published for Bs-DHQase, it might be speculated that it is inactive owing to the lack of a catalytic base able to abstract a proton from the C2 of the substrate. In all other enzymes, the active-site Tyr23 is part of the flexible loop, which is either not visible or has different conformations in different subunits of the present structure.

In the active site, one of the structured water molecules is found in the same position in all known structures of type II DHQases. This water molecule has been proposed to play a role in the protonation of the enol intermediate and as such is of crucial importance for catalysis. Another water molecule in this region is bound to the main-chain O atom of Pro10 and is located in a position that could correspond to the carbonyl O atom of the substrate.

One of the sulfate atoms in the active site occupies the carboxylate-binding site of the dehydroshikimate in the mutant structure and of the 2,3-anhydroquinic acid in the native Sc-DHQase structure. It is ligated to Ser102 and Asn74, which are both conserved in the type II DHQases of known structure. We compared the sulfate position in the present

structure with the position of the sulfate and the phosphate in the Mt-DHQase and Sc-DHQase structures, respectively. The three anions occupy practically the same position. Nevertheless, it is surprising that the sulfate binds in a conformation that is similar to the phosphate in Sc-DHQase and that differs from the sulfate in Mt-DHQase (Fig. 6).

It has been speculated that a cavity behind His105 in Sc-DHQase may play a role in the release of the product of the reaction (Roszak *et al.*, 2002). In the present structure, this cavity is partially filled with Cys125, which still has enough space to take up three conformations in all 12 monomers. In Bs-DHQase this cavity is reduced by the larger side chains Leu, Tyr and Gln that respectively replace Val104, His105 and Val123 in Ap-DHQase. However, the biological significance of this finding is not clear since Bs-DHQase is likely to be inactive as discussed earlier.

3.6.4. The non-specific ligands. Firstly, we compared the molecules located on the four threefold axes of the molecule. In all the Sc-DHQase structures, a Tris molecule is located on each axis in the same position as in our structure. The sulfate molecule on each threefold axis is also present in the sulfate-liganded Mt-DHQase structure. It is located at a distance of 3.2 Å from the sulfate in our structure on the opposite site of the Tris molecule. In the phosphate-liganded Sc-DHQase structure, on the other hand, the phosphate occupies the same position as the sulfate in our structure. Nevertheless, the conformation is different, with the O atom of the anion pointing in the opposite direction: it points towards the Tris molecule in the Sc-DHQase structure and away from it in the present structure (Fig. 7). This O atom is surrounded by a symmetrical tetrahedral cluster of conserved waters. In our structure, this highly symmetric network makes contacts with

the side chain of Glu55. The equivalent amino acid in Sc-DHQase is a glycine. The location of the side chain of His53 (Ap numbering) in Sc-DHQase, which is a glycine in Ap-DHQase, precludes the formation of this network in Sc-DHQase. Nevertheless, a similar network is found on the opposite side, bonded to the O atom on the threefold axis of the phosphate molecule. From this network the three waters symmetrically positioned around the threefold axis are also found in the present Ap-DHQase structure. None of the other DHQase structures has an anion positioned on the threefold axes.

A further comparison of the other sulfate positions in our structure with those in the sulfate-liganded Mt-DHQase structure shows that the sulfate at the entrance of the active site, which is liganded to the flexible loop, is in a different position in the two structures (2.3 Å apart) owing to the different conformation of the residues at the base of the flexible loop (Fig. 6). The surface sulfates in the two structures are located in different regions owing to the different composition of the neighbouring amino acids.

4. Conclusion

We have presented the crystal structure of DHQase from *A. pleuropneumoniae*. The use of the counter-diffusion technique increased the diffraction limit of the crystals from 3 to 1.7 Å. This high resolution allowed us to obtain a detailed view of the active site in the uncomplexed form.

It is a common feature of enzymes to have an active site that is partially rigid and partially flexible. This is particularly common for transferases. The X-ray structures of such enzymes show that in many cases the enzyme can exist in open

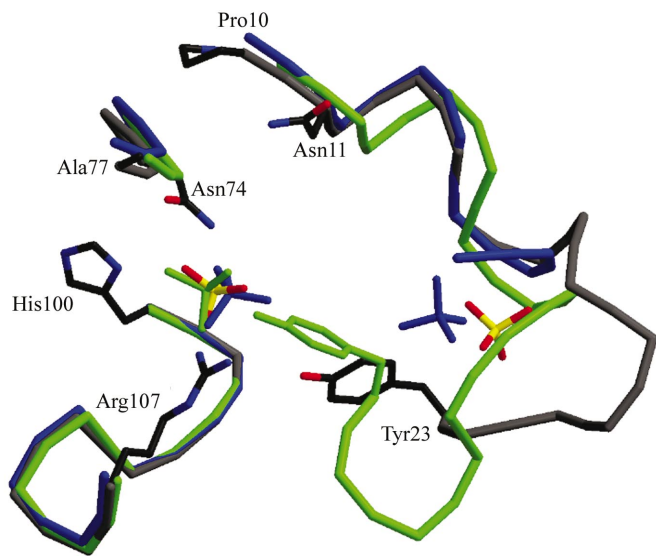


Figure 6

Superposition of the active site and the active-site anions of the present sulfate-liganded Ap-DHQase (main chain, grey; sulfates, yellow), the phosphate-liganded Sc-DHQase (green; PDB code 1d0i) and the sulfate-liganded Mt-DHQase (blue; PDB code 1h05). The active-site residues of the Ap-DHQase structure are shown, as well as Tyr23 from Sc-DHQase in the closed conformation.

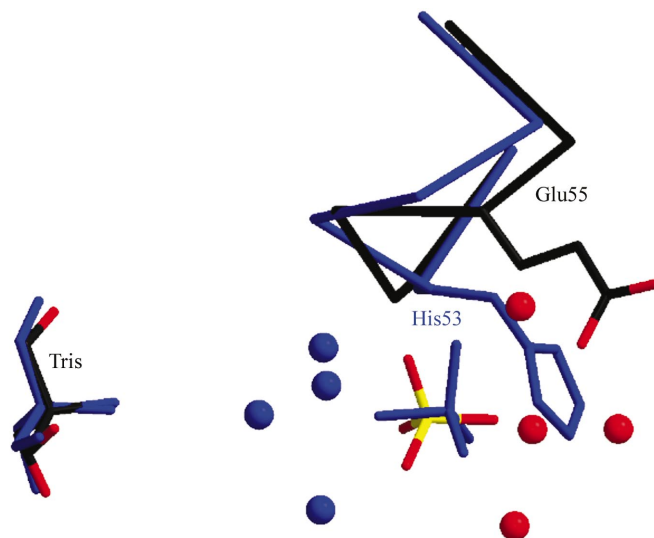


Figure 7

View of the water clusters (red balls) and the anions along the threefold axis after superposition of the Ap-DHQase (main chain and Tris in black, sulfate in yellow and red) and the phosphate-liganded Sc-DHQase (main chain, Tris and phosphate in blue; PDB code 1d0i). The side chain of Glu55 from Ap-DHQase and of His53 (Ap numbering) from Sc-DHQase are also shown.

or closed conformations, where ligand binding induces the closed conformation. The open form facilitates ligand binding and the closed form prepares, controls and protects the reaction intermediates. In the cases where information of the loop dynamics is available, the loop motion is either rate-limiting or closely coordinated with the chemical steps of the reaction (Desamero *et al.*, 2003; Rozovsky *et al.*, 2001; Wang *et al.*, 1999). In Ap DHQase the rigid part of the active site forms a pocket. The flexible loop is either completely disordered or present in one of two sulfate-bound conformations. The deprotonated form of Tyr23, which belongs to the loop, has been proposed to act as a general base, but in none of the monomers does it take up a position similar to the closed inhibitor-bound form of the enzyme (Roszak *et al.*, 2002). Thus, the different monomers in the present structure represent different conformations of the open form of Ap-DHQase.

In the present work, we have shown that a sulfate ion binds in the active site of Ap-DHQase in the position of the phosphate in the phosphate-liganded Sc-DHQase. A high percentage of known proteins bind or process compounds possessing phosphoryl groups. Because sulfates have similar physical properties and are of similar size, sulfate ions can bind in most phosphate-binding sites (except the HPO_4^{2-} site, which binds phosphate in its monobasic form and has a high degree of specificity). However, sulfate is less frequently implicated in a specific functional role than phosphate, resulting in a more limited biochemical interest in sulfate-binding sites, which are often located on the surface of a protein.

Our results suggest that a sulfate ion can bind in the active site with the flexible loop in the open conformation. The presence of the sulfate and the contacts it makes are not sufficient to stabilize the closed conformation as the phosphate does in the Sc-DHQase structure. An open conformation of an active-site flexible loop in the presence of a sulfate and a closed conformation in the presence of a phosphate have previously been observed in triosephosphate isomerase (TIM; Wierenga *et al.*, 1992); in the presence of sulfate, crystal-packing contacts were necessary in order to stabilize the closed conformation of this loop. For the enzyme TIM, evidence was provided for the view that loop motion and product release are concerted and might be one of the rate-limiting steps in catalysis. For the DHQase no conclusive evidence is yet available to support this idea.

LAG acknowledges CONACyT for financial support. We thank Professor Kabsch for providing XDS free of charge. We acknowledge the support of the European Space Agency and the European Community Action to Research Infrastructure Action of the Improving Human Potential Programme to the EMBL Hamburg Outstation, contract No. HPRI-CT-1999-00017. We thank Olivier Minster (ESA) for his support of

space science. The authors acknowledge the excellent work of Dr Eva Mañas in managing the logistics concerning the space mission. We thank Viscount Dirk Frimout for his support for space crystallization experiments.

References

- Aiyar, A. (2000). *Methods Mol. Biol.* **132**, 221–241.
- Bosse, J. T., Janson, H., Sheehan, B. J., Beddek, A. J., Rycroft, A. N., Simon, K. J. & Langford, P. R. (2002). *Microbes Infect.* **4**, 225–235.
- Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst. D* **54**, 905–921.
- Coggins, J. R., Abell, C., Evans, L. B., Frederickson, M., Robinson, D. A., Roszak, A. W. & Laphorn, A. P. (2003). *Biochem. Soc. Trans.* **31**, 548–552.
- Desamero, R., Rozovsky, S., Zhadin, N., McDermott, A. & Callender, R. (2003). *Biochemistry*, **42**, 2941–2951.
- Engh, R. A. & Huber, R. (1991). *Acta Cryst. A* **47**, 392–400.
- Evans, L. D., Roszak, A. W., Noble, L. J., Robinson, D. A., Chalk, P. A., Matthews, J. L., Coggins, J. R., Price, N. C. & Laphorn, A. J. (2002). *FEBS Lett.* **530**, 24–30.
- Garcia-Ruiz, J. M., Gonzalez-Ramirez, L. A., Gavira, J. A. & Otalora, F. (2002). *Acta Cryst. D* **58**, 1638–1642.
- Garcia-Ruiz, J. M., Moreno, A., Otalora, F., Viedma, C., Rondon, D. & Zautscher, F. (1998). *J. Chem. Edo.* **75**, 442–446.
- Garcia-Ruiz, J. M., Novella, M. L., Moreno, R. & Gavira, J. A. (2001). *J. Cryst. Growth*, **232**, 165–172.
- Garcia-Ruiz, J. M., Otalora, F., Novella, M. L., Gavira, J. A., Sauter, C. & Vidal, O. (2001). *J. Cryst. Growth*, **232**, 149–155.
- Gourley, D. G., Shrive, A. K., Polikarpov, I., Krell, T., Coggins, J. R., Hawkins, A. R., Isaacs, N. W. & Sawyer, L. (1999). *Nature Struct. Biol.* **6**, 521–525.
- Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
- Kleywegt, G. J. & Jones, T. A. (1994). *CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **31**, 9–14.
- Kraulis, P. (1991). *J. Appl. Cryst.* **24**, 946–950.
- Lalonde, G., O'Hanley, P. D., Stocker, B. A. & Denich, K. T. (1994). *Mol. Microbiol.* **11**, 273–280.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Lee, B. I., Kwak, J. E. & Suh, S. W. (2003). *Proteins*, **51**, 616–617.
- Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.
- Navaza, J. (1994). *Acta Cryst. A* **50**, 157–163.
- Navaza, J. (2001). *Acta Cryst. D* **57**, 1367–1372.
- Roszak, A. W., Robinson, D. A., Krell, T., Hunter, I. S., Fredrickson, M., Abell, C., Coggins, J. R. & Laphorn, A. J. (2002). *Structure*, **10**, 493–503.
- Roussel, A. & Cambillau, C. (1991). *Silicon Graphics Geometry Partners Directory*, p. 81. Mountain View, CA, USA: Silicon Graphics.
- Rozovsky, S., Jogl, G., Tong, L. & McDermott, A. E. (2001). *J. Mol. Biol.* **310**, 271–280.
- Rycroft, A. N. & Garside, L. H. (2000). *Vet. J.* **159**, 18–36.
- Wang, G. P., Cahill, S. M., Liu, X., Girvin, M. E. & Grubmeyer, C. (1999). *Biochemistry*, **38**, 284–295.
- Wierenga, R. K., Noble, M. E. & Davenport, R. C. (1992). *J. Mol. Biol.* **224**, 1115–1126.